Feedforward neural networks for very short term wind speed forecasting

Citation for published version:
Tagliaferri, F & Viola, IM 2012, 'Feedforward neural networks for very short term wind speed forecasting' Paper presented at Marine and Offshore Renewable Energy, London, United Kingdom, 26/09/12 - 27/09/12, pp. 85-88.

FEEDFORWARD NEURAL NETWORKS FOR VERY SHORT TERM WIND SPEED FORECASTING

F Tagliaferri and **I M Viola,** Yacht and Superyacht Research Group, School of Marine Science and Technology, Newcastle University, UK

SUMMARY

Since 2007 wind has become the major source of renewable energy in the UK. Moreover, increasing oil costs are driving researchers in the marine transport field to develop innovating wind ships. In order for wind power to be effectively and efficiently exploited, reliable forecasts on wind speed are needed. These will allow saving curtailments costs, improving safety, reducing damages due to extreme weather conditions, etc. Also, short and very short wind forecasts are critical for energy trading. In this study we present a short-term wind forecast based on artificial neural networks, which are mathematical structures able to model complex non-linear systems. In particular, we used a multilayer perceptron that predicts future wind speed values given the past and current recorded values. Data sampled every ten minutes was used to forecast up to one hour ahead, with an uncertainty ranging from 5%, for ten minutes ahead forecast, to 21%, for one hour ahead forecast.

1. INTRODUCTION

The 2011 annual wind report of the World Wind Energy Association (WWEA) [1] stated that, in 2011, the worldwide wind energy production increased by 40,053 MW, reaching a total capacity of 237,016 MW, which corresponds to the biggest increase in history, but also corresponding to a growth rate of 20,3% which is the lowest in the last decade The European Union has approved in 2008 the climate and energy package (known as 20/20/20 strategy) whose aim, by the year 2020, is to reduce greenhouse gas emissions by 20%, to establish a 20% share for renewable energy, and to improve energy efficiency by 20% [2]. In order to meet this challenging target, deployment and operation of wind energy device need a significant step change. In particular, there is an unmet need for wind forecasting. More specifically, while long term forecasts (up to five days) allow decisions for deployment and maintenance of the structure; short (up to 24 hours) and very short (up to 1-2 hours) are critical for energy trading [3,4].

There are two possible approaches to wind power forecasting: one is to predict wind velocity and then use a power curve to convert it to energy production; another one is to predict directly the wind energy. The choice between these two approaches needs to take into account that also the relationship between wind speed and wind energy production can have a stochastic nature [5] or being highly nonlinear [6].

In this work we present a model for the prediction of wind velocity. Models for predicting wind speed can be numerical or statistical. Numerical model are based on mathematical fluid mechanics models and have dominated the literature until the last decade (for instance, [7,8]). Statistical methods are based on past observation of the wind behaviour at one or more locations and have been found to perform better than numerical models on very-short-term forecasts [9,10]. In this paper we present a statistical model based on Artificial Neural Networks (ANN) using past recorded values of wind speed to predict the future values.

ANN have been successfully used to attack a wide range of problems such as, for instance, speech recognition [11], image classification [12], function approximation [13] and financial forecasting [14]. The present study uses ANN to predict the future wind velocity based on several successive sets of velocity measurements taken at a single location.

Differently from earlier studies using ANNs for wind forecasting [15], where temperatures, humidity and pressures were also input to the model, our model uses only the recent past velocities allowing a fast training of the network.

2. METHOD

ANN are inspired by the functioning of the biological neural networks in the brains of humans and animals, and their peculiarity is the possibility of emulating the human process of learning from experience. The constitutive unit of a neural network is a neuron, which is a singular processing unit that takes several inputs x_i originating from other neurons, and produces an output that is then transmitted to other neurons. The mathematical representation of the structure of a neuron is shown in Figure 1. A neuron itself can be broken down into the following components:

- A set of connecting links, called synapses, where the $i-th$ synapses is characterized by a weight w_i (synaptic weights);
- An adder within the neuron that performs a sum of the inputs weighted by the corresponding synapses;
- An activation function φ, which transforms the sum computed by the adder into the neuron output. If the activation function is linear, a neuron results in a linear combination of the input values, while non-linear activation functions (generally sigmoid functions) allow for the modelling of non-linear problems.

Therefore, a neuron can mathematically be described by Equation (1):

$$y = \varphi(\zeta); \;\; \zeta = \sum_{i=1}^{u} w_i x_i \tag{1}$$

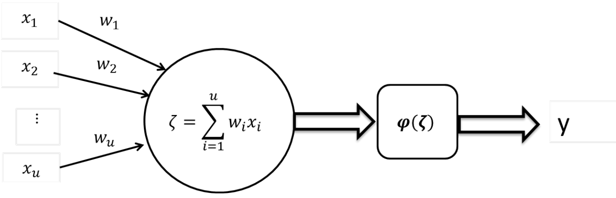

Figure 1. Structure of a generic neuron.

Neurons are assembled together into an integrated structure that depends on the kind of problem that the network has to solve. A structure that has been successfully used in function approximation is the so-called feed-forward multi-layer perceptron, characterized by the organization in subsequent layers of the neurons, as can be seen in the schematic example in Figure 2.
The learning process involves the continuous modification of the synaptic weights and it is based on the principle of iterative error-correction. The synaptic weights of the various neurons are initialized to random values, then a training set of input and output data is presented to the network.
For each input vector the initially generated output vector is compared with the known true output vector. The synaptic weights of the output layer are then modified by adding a factor that is proportional to the current assessed error and to a learning rate, and those corrections are extended to all of the weights in the network through a back-propagation process. This operation is iterated until successive changes in the synaptic weights are smaller than a given value, or when the errors begin to increase. For further details on training algorithms and validation processes see [16].

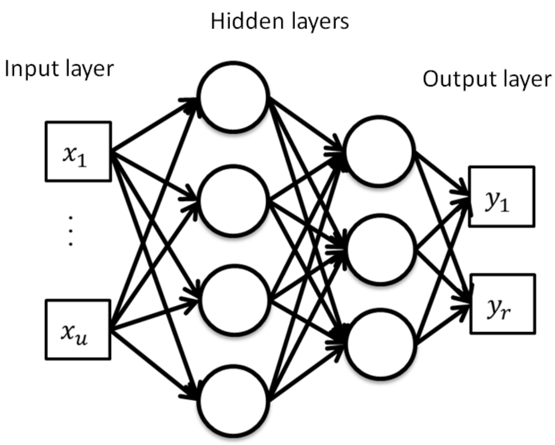

Figure 2: Schematic diagram of a multi-layer feed-forward perceptron.

In this work we use a multilayer perceptron to perform a forecast for wind speed based on past values. A time series approach is used: we assume that there exists a function f such that:

$$s(t_{k+1}) = f\big(s(t_k), s(t_{k-1}), \ldots, s(t_{k-m})\big) \tag{2}$$

where $s(t)$ represents the wind speed at time t. In the present study the time values for the sampling $t_k$are taken at a distance of 10 minutes.
Therefore, we are looking for a way to express the next future value for the wind speed as a function of a vector of past values. The ANN is trained in order to model the function f. An input vector, consisting of consecutive measured values for the wind speed, is used as input, while future values are used as training outputs for the network.

3. RESULTS

A multi-layer perceptron was used to perform a forecast on ten-minutes wind speed measurements. The data set, provided by the National Climate Database and available online [17], was made of 4000 consecutive measured values; 80% of the values were used for the training, while the remaining 20% were used for testing the performance of the trained ANN.
Different networks were trained in order to identify the best structure to perform the forecast.
Single layer perceptrons can be used only to model linearly separable problems [16], and tests confirmed this limitation as it was observed that the performance of such a network was highly dependent on the initialization of the synaptic weights. Conversely, two hidden layers allow modelling non-linear functions, such as f in Eq. (2). Increasing both the number of neurons per layer and the size of the input vector - increasing m in Eq. (2) - led the average error to decrease until optimum values are reached. Then a further increase in neurons led the performance to decrease again. Both an excessive number of neurons and a too large size of the input vector led the performance to decrease because the number of parameters to be optimised increases and the training becomes inefficient. For instance, for our data set, the best performance was achieved with two hidden layers with 18 and 15 neurons, respectively, and with an input vector made of eight consecutive wind speed measurements - $m = 7$ in Eq. (2). The single output vector predicted the ninth consecutive value.
Having two different non-linear activation functions for the two layers increases the learnability when dealing with highly non-linear models [16]. For instance, we used a log-sigmoid and an hyperbolic tangent activation function for the first and second layer, respectively.
The training is performed with the Levenberg-Marquardt back-propagation algorithm, which is known to be efficient for networks with less than 100 neurons [18].
Figure 3 shows the comparison between wind speed values registered by the station and wind speed forecast

by the ANN. The signal corresponding to the wind speed is highly oscillating and no qualitative general trend can be extrapolated. However, the ANN is able to perform a forecast with an uncertainty of 5% at 95% confidence level (meaning that, in 95% of the cases, the error is less than 5% of the average the wind speed).

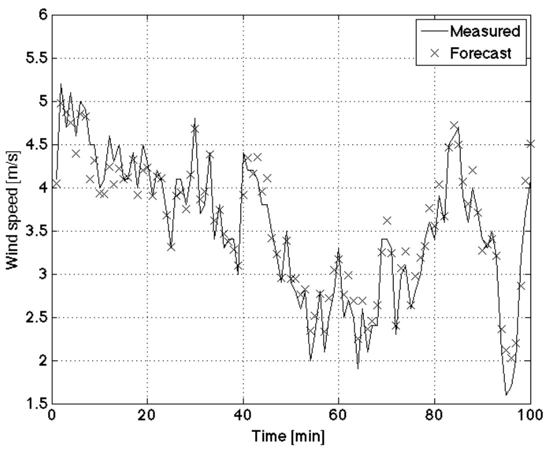

Figure 3: Measured and forecast wind speed.

The same method with a different training can be used in order to predict the successive-step-ahead wind speed value. Therefore, being $s(t_k)$ the most recent measured wind speed known by the ANN, the input vectors are still of the form $[s(t_k), s(t_{k-1}), \dots, s(t_{k-m})]$, while the output is $s(t_{k+j})$, with j varying from 0 to 6. We use again $m = 7$. The forecast uncertainty increases with j. In particular, Table 1 shows the uncertainty for the different minutes ahead forecasts. Ten minutes ahead is computed with $j = 1$, and one hour ahead is computed with $j = 6$.

Table 1: Uncertainty for different minutes ahead forecasts.

Minutes ahead [min]	10	20	30	40	50	60
U [%]	5	7	11	15	18	21

4. CONCLUSIONS

Reliable short and very short wind energy production forecasts are necessary for energy trading. In order to obtain a good forecast it is possible to transform a forecast for wind speed into a forecast for wind energy production for a wind farm. In this paper, we presented a model for short-term wind forecast based on artificial neural networks. The model uses ten-minutes wind measurement and, taking as input eight consecutive wind values, predicts the wind speed for the next values up to one hour ahead. The uncertainty increases almost linearly with the time distance between the last input values and the desired output. The model allows forecasting up to one hour ahead with an uncertainty of 21%.

5. REFERENCES

1. GSANGER, S., PITTELOUD J. D., 'WWEA Report 2011', 2011.

2. EUROPE Press Release: "Climate change: Commission welcomes final adoption and energy package", available at the web page http://europa.eu/rapid/pressReleasesAction.do?reference=IP/08/1998.

3. PARKES, J., WASEY, J., TINDAL, A., MUNOZ, L., "Wind Energy Trading Benefits Through Short Term Forecasting" *European Wind Energy Conference technical Papers,* 2006.

4. POTTER, C. W., "Very short-term Wind Forecasting for Tasmanian Power Generation" *IEEE Transactions on Power Systems,* 2006.

5. JEON, J., TAYLOR, J. W., 'Using Conditional Kernel Density Estimation for Wind Power Density Forecasting', *Journal of the American Statistical Association,* 2012.

6. SANCHEZ, I., "Short-Term Prediction of Wind Energy Production", *International Journal of Forecasting, 22, 43-56, 2006.*

7. COTTON W. R., PIELKE R. A., WALKO R. L., LISTON, G. E., TREMBACK, C. J., JIANG, H., MCANELLY, R. L., HARRINGTON, J. Y., NICHOLLS, M. E., CARRIO, G. G., MCFADDEN, J. P., "RAMS 2001: Current status and future directions" *Meteorology and Atmospheric Physics, 82 (1-4), 5-29*, 2003.

8. JACOB D., PODZUN, R., "Sensitivity studies with the regional Climate Model REMO", *Meteorology and Atmospheric Physics 63(1-2), 119-129,* 1997.

9. LI, S., "Using neural networks to estimate wind turbine power generation," *IEEE Trans. Energy Convers., 16(3), 276–282,* 2001.

10. ALEXIADIS M. C., DOKOUPULOS P. S., SAHSAMANOLOGOU H. C., "Wind Speed and Power Forecasting Based on Spatial Correlation Models," *IEEE Trans. Energy Convers., 14(3), 836–842,* 1999.

11. MORGAN N., BOURLARD H. A. (1995), "Neural Networks for Statistical Recognition of Continuous Speech", *Proceedings of the IEEE, 83 (5), 742-770.*

12. LAWRENCE S., GILES C.L., TSOI A. C., BACK A.D., "Face Recognition: A convolutional Neural-Network Approach",

IEEE Transactions on Neural Networks, 8(1), 98-113 1997.

13. POGGIO T, GIROSI F., "Networks for Approximation and Learning", *Proceedings of the IEEE, 78(9), 1481-1497*, 1990.

14. KAASTRA I., BOYD M., "Designing a Neural Network for Forecasting Financial and Economic Time Series", *Neurocomputing, 10(3), 215-236*, 1996.

15. MORE A., DEO M. C., "Forecasting Wind with Neural Networks", *Marine Structures, 16(1), 36-49,* 2003.

16. HAYKYN S., "Neural Networks, A Comprehensive Foundation", *Prentice Hall PTR Upper Saddle River NJ USA,* 1999.

17. NIWA, National Climate Database, http://cliflo.niwa.co.nz/, visited on 10th May 2012

18. HAGAN, M., MENHAJ, M. B., "Training Feedforward Networks with the Marquard Algorithm", *IEEE Transactions on Neural Networks, 5(6), 989-993,* 1994.

6. AUTHORS BIOGRAPHY

Francesca Tagliaferri is a PhD student at Newcastle University and member of the Yacht and Superyacht Research Group. She holds a Masters degree with Honours in mathematics and her PhD project aims at developing a navigation software for yacht races under uncertain weather conditions.

Dr. Ignazio Maria Viola is Lecturer in Naval Architecture at the Newcastle University and group leader of the Yacht and Superyacht Research Group. His specialist skills and experience are in numerical and experimental incompressible viscous fluid dynamics. In particular, his main research focus is the hydro and aero dynamics of yachts and superyachts, while his research portfolio also includes the fluid dynamics of renewable energy devices, trains, cars, long-span bridges and tall buildings. In 2008 he successfully completed a CFD simulation with a record breaking grid of more than one billion cells and in 2010 he performed the world's first direct pressure measurements on a full-scale spinnaker. Ignazio has more than 40 peer reviewed scientific publications on the fluid dynamics of yachts and super yachts.